**San Joaquin River Fall-run Chinook Salmon Population Model**

**External Scientific Review Form**

**Reviewer:** #4

**Review:**

1. **Problem/Goals.** Is the problem that the project is designed to address adequately described? Are the goals, objectives and hypotheses clearly stated and internally consistent?

The problem is well defined, and the goals are clearly stated.

2. **Approach.** Is the approach well designed and appropriate for meeting the objectives of the project as described in the proposal?

See "Is the model adequate?" section below.

3. **Feasibility.** Is the approach fully documented and technically feasible? What is the likelihood of success? Is the scale of the project consistent with the objectives?

Not applicable

4. **Project Performance Evaluation Plan.** Will a monitoring plan be developed to document changes in the restored habitat over time and the response of salmonids and/or riparian vegetation to the restoration in a scientifically rigorous manner?

Not applicable

5. **Expected Products/Outcomes.** Are products of value likely from the project?

Not applicable

**Additional Questions:**

**General:**
The purpose of the model is to develop spring flow magnitude, duration, and frequency instream flow levels into the South Delta to adequately protect, and restore, fall-run Chinook salmon in the San Joaquin River basin. To accomplish this objective, please address the topics listed below for these questions:

Is the model adequate?

Whether the model is adequate or not depends on the questions asked and how the results are interpreted. I think that the model is adequate to conclude that higher flows released over a broader time window later in the season would benefit the salmon. I think the model is inadequate if the answers need to be much more precise than that. Maybe one can deduce the rough magnitudes of the changes needed but not in much detail. I do not think the model, and the way it was applied, can produce accurate and precise enough predictions to make specific recommendations such as stated in the Executive Summary. I do not think the model contains enough biology to be able to quantitatively distinguish the long-term, multi-year effects of 5,000 versus 10,000 cfs or between the effects of a 30 day versus 45 day window. Further, when much larger differences are simulated, the predicted response in escapement seems, at times, to be unrealistic (i.e., >100% increase). This is likely due to reliance on mostly linear statistical relationships. What is critical is how the results are interpreted.

I could go into questions I have about specific relationships but you get the idea.

I will add up front in my review that I believe the model was carefully thought out and the data carefully analyzed. The description of the model is generally thorough and report is generally well written and understandable.

If not, how can model be improved?

In my opinion, the weaknesses of the model stem from its over-reliance on statistical correlations, and the appearance of ad-hoc decisions as to which statistical relationships are strong versus weak and thus included or excluded from the model. I do not agree with the principle assumptions of the model which removed ocean harvest, exports, and density-dependence from further consideration. I am also familiar with some of the more critical datasets used to derive the relationships, and have participated in debates about the interpretation and validity of the correlation relationships used to analyze the data, and which now appear in the model. Some of these relationships rely heavily on a few points which control the slope of the relationship. One may plug in x-values within the range of the data but the resulting y-value is still very sensitive to the few extreme points that control the slope.

1. Foundation (justification) – The reasons for developing the model is well stated. The devil is in the details of how should the model be developed. I am sympathetic with the author's situation: the model must be defensible which pushes one to statistical relationships. Yet, sometimes the questions posed are not well suited to a regression-based population dynamics model. I think this is one such situation, especially since the desired answers need to be relatively precise. I think the model could be improved by the development of a parallel version that incorporates more biology and less reliance on whether statistical correlations of field data were significant or not. The two models would bound the answers.

2. Logic – The author lays out a very logic approach to developing the model. I understand his logic; I just disagree on philosophical grounds. But there is a saying that you put two modelers in a room with a problem and they come out with 3 models.

3. Numeric representations – I get a little nervous with models done in excel. The user interface is very nice but there can be problems with excel and numerical calculations. I would like to see some confirmation that excel, and the visual basic or however the model is represented in excel, is performing the computations correctly.

4. Application and reliability – The author does a nice job using a systematic approach to exploring how the magnitude and timing of flow would affect salmon escapement. I would likely use a similar simulation experimental approach, but with a different model. The author acknowledges that the computed confidence intervals are really not the appropriate variability around model predictions. This relates to the reliability issue. One can use a Monte Carlo approach or bootstrapping to derive more appropriate error bars on the final predictions. The treatment of water-type years is an excellent first step towards increasing reliability, but only addresses one of many possible sources of uncertainty and stochasticity.

5. Conclusions -- The actual conclusions in the report are reasonable; the specific recommendation in the Executive Summary is OK as long as it viewed and interpreted correctly. In my opinion, the highly quantitative model results should be viewed as qualitative or semi-quantitative predictions. The model results should be taken with other sources of evidence to determine whether it is worth changing VAMP. I personally think the results are probably pretty good, but if I am asked to place all of confidence on the model results then I back off somewhat.

6. Calibration and validation – The calibration and validation is weak in the present application of the model. It was not clear which regression coefficients (which are also model parameters) were adjusted by the author, and whether this was done in a systematic way or not. I wonder why the author did not simply use solver in excel and optimize the calibration. There is no model validation.

7. Documentation -- The documentation is pretty good. It could be improved by a table or appendix that actually lists the equations of the model and the order of computation. More information on the calibration method and which coefficients were changed and by how much in the final calibrated version from the data-derived estimates would help. It might be interesting to see if the adjusted coefficients result in relationships that still fit the data upon which they first estimated from.

8. Testing (i.e. what monitoring could occur to validate or reject model predictions) – I think the author tried to use the available information. Indeed, many of the correlations and arguments in the report approach become circular and convoluted because the same data seems to be used in multiple ways. The population model cannot really be tested at the level that would be ideal (i.e., the long-term population level). Perhaps one can use a more detailed model, less constrained by significant statistical relationships existing or not, to see if the same results would be predicted.

**Specific:**

**<u>Hydrology</u>**

1. Are the methods used in the Model (including Model Report) relating to flow sufficiently documented? If not, what improvements can be made to improve documentation?

I think adding a background section that explains exports, delta inflow, etc. in the context of the salmon life history would help. Nothing extensive, but an overview of the hydrology overlaid on the life cycle. Also, a graphical presentation the features of the different water-type years would be helpful.

2. What is the best metric (i.e. arithmetic mean, geometric mean, transformed data etc) that can be developed to adequately capture the variability in spring flow (i.e. magnitude and duration) on an intra-annual basis?

I am not sure.

3. What improvements to hydrologic data utilization can be made to enhance model prediction performance reliability?

Model simulations that use a variety of sequences of year-types would help model reliability, especially when the model is used to make forecasts. Also, what is the variability among years of the same water-year type? Do the patterns of water flows overlap among designations, or are they unique? I recall that specific years can look like one type in spring and another type in the summer. How is this dealt with? Perhaps using water-year types on a seasonal basis might help with reliability? At the other end of the scale, there is no mention of how climate change might affect the water flows. Will we expect to see more high-flow years occurring together, etc.?

4. Is there evidence of auto-correlation in flow calculations? If so, what is the affect? Does it need to be removed to improve model prediction (flow determination) reliability? If so, how can it be removed?

I can think of two levels of auto-correlation: day-to-day and year-to-year. I think the day-to-day is OK. The year-to-year is more problematic and more interesting. I suggested in (3) above that more investigation of how the sequences of year types affects predictions. This is a direct approach to adjusting the inter-annual correlation (either making it higher or lower).

5. Are there additional flow metrics, parameters, logic etc. that should be incorporated into model logic and function? If so, what are they and how can they be assimilated into the model (reference to logic and function)?

I am not sure about the question but I will use this opportunity to suggest that I would also look at how salmon habitat itself (without regression relationship to convert it to escapement) would change under the different scenarios. Before one even gets to salmon numbers, one would know how the physical habitat will change.

**Biology**

1. Are the methods used in the Model (including Model Report) relating to fish abundance and/or production sufficiently documented? If not, what improvements can be made to improve documentation?

The entire model is series of linked regression models. I think a table or appendix that shows the exact equations used would be helpful in terms of documentation. There are some details that would help, such as how the final escapement is determined from the multiple years of contribution. I can tell petty much what the model is by the report, but I am very familiar with population models.

2. What improvements to fish data utilization can be made to enhance model prediction performance reliability?

We know a lot more about salmon reproduction, growth, mortality, and movement than is implied by this model. I think the fish data can be further utilized to put more biology into the model. Perhaps a complementary model, less constrained by statistical significance, would help. Also, I think the data can be mined more for factors affecting interannual variation in salmon.

4. Is there a way to improve how the model performs fish abundance prediction calculations and/or processing of fishery data?

The model can be calibrated using formal minimization methods, and uncertainty analysis can be used to show probability distributions of predictions.

5. Is there evidence of auto-correlation in fish related calculations? If so, what is the affect? Does it need to be removed to improve model prediction (flow determination) reliability? If so, how can it be removed?

Their must and should be auto-correlation in the fish calculations. Escapement is comprised of multiple year-classes. Auto-correlation in the fish can also arise from auto-correlation in the hydrology. I would not want it to be removed. I might like to see how important the auto-correlation was so that I can judge how the historical time series (with its built-in auto-correlation) might have affected the predictions. One way to asses this is remove it completely (run with a random sequence of year types) and to manipulate the sequence of year types. This would be more as sensitivity analysis and would be helpful for interpreting the results based on the historical time series.

6. Does justification exist to include additional fish metrics, parameters, logic etc. in model logic and function (i.e. ocean harvest and/or Delta export entrainment)? If so, what are they and how can they be numerically assimilated into the model (reference to logic and function)?

I think export, ocean harvest, density-dependence, and stochasticity (real year-to-year and with-in year variability as it occurs in nature) should be included in the model.

7. How can model representation of hatchery production, and underlying model logic, be improved upon?

As noted by the author, how hatchery salmon will react to flow and temperature are unknown. Will hatchery fish follow the same relationships to flow as was (sometimes weakly) estimated for natural fish? The way hatcheries were included was consistent with the general modeling approach used with natural fish.

8. Currently the model predicts a constant ocean survival rate (i.e. relationship between cohort abundance and Chipps Island abundance is constant). Is there a need to make this relationship variable? If so, how can this be numerically accomplished in model performance?

Making ocean survival a constant is OK as a first step. All model predictions are then conditioned on this constant ocean mortality rate. I would make ocean mortality variable in the model.

9. The model currently uses an adult replacement ratio of 1:1 as a numerically identified population health barometer. Is there a need to refine this ratio? What additional population parameter(s) could be incorporated into model logic and function?

An adult replacement ratio is a reasonable health metric. Whether it should be 1:1 or something else depends on the management objective and whether 1:1 will get you recovery in the time frame needed. There are other metrics, but their use is limited by the structure of the model and what it can predict. I think looking at what the adult replacement ratio metric means in terms of survival in other life stages would be helpful. Can the conditions that resulted in 1:1 ratio be expected to produce the required reproduction or survival in other life stages through the life cycle?

**<u>Statistics</u>**

1. Currently the model uses liner relationships between flow and fish production because this relationship provides the strongest correlation value. Is it necessary to include a model toggle switch, model logic, and mathematical functions, that allow users the option to test a variety of non-linear relationships between flow and fish survival and/or production upon model results?

The short answer is yes – but even including options on linear relationships would not address my major concern of the general modeling approach that relies on a series of linked correlations. Linear relationships just make me much more nervous. We know from experience that correlation-based population model of fish do not work very well for long-term predictions. They will evidently fall apart because one or more of the correlations falls apart. This does not mean that useful information cannot be obtained from a correlation-based modeling effort. But it means that appropriate caution should be used in interpreting the results. The various correlations are based on different (often incompletely overlapping) years, with the associated differences in environmental and biological conditions.

2. What is the statistical reliability of model out-put given that model predictions propagate? How can model reliability be improved?

The reliability of the final model predictions is low in absolute terms (i.e., actual escapement). This is typical for population dynamics models, both correlation-based and process-based. More accurate confidence intervals would undoubtedly show that most of the different scenarios generate overlapping predictions of escapement. Also, the dismissal of variation in ocean harvest and the dismissal of other factors completely make the model predictions not accurate at the absolute level. The model is best used to generate relative changes, which the author highlights in the results tables (percent change from historical). There are techniques available for propagating uncertainty and stochasticity through this type of model. In fact, I think there is an add-in to excel (something called crystal ball) that would allow one to use Monte Carlo techniques to analyze the uncertainty in the model.

3. Is collinearity present in model logic and/or computation, and what influence does it have upon model results? If present how can it be removed?

I think there is the potential for a great deal of collinearity in the model because the same datasets seemed to be used in multiple ways. For example, lagged escapement is used to derive the arrivals at Mossdale to begin the model, and of course, escapement is the primary prediction variable. There are many examples of this. Another source of collinearity is the multiple uses of the same data. For example, the x-axis for the same data was expressed as export ratio, exports, and Vernalis flow plus exports. All of the resulting relationships were questionable (high scatter, outliers, cluster of influence points) relationships and the conclusion of no need to include exports in the model was a major assumption. The same approach was used to dismiss density-dependence from consideration. The best way to remove collinearity is to build the model differently.

4. In some cases, model predictions for salmon production occur outside the empirical data set range used to develop the regression. What limitations in model reliability result?

This focus on “outside-of-the-range” is valid but I am concerned that this implies that predictions within the range are therefore considered OK. Outside-of-the-range predictions are clearly a problem. Inside-the-range predictions are also a problem, as the model uses the slopes of relationships, which can be greatly influenced by a few points. The model can be used to generate qualitative or semi-quantitative predictions but not quantitative predictions. Also, the predictions presume all other factors remain constant, so the model is less reliable in making predictions for say the next 3 years. Over may years, one could hope that the other factor average out in their effects.

5. Presently smolt survival has a statistically significant regression correlation with Delta inflow level (i.e. less than 7,000). No statistically significant regression correlation for juvenile smolt survival and Delta export level exists. However when inflow to export ratio is regressed against flow survival, a moderate regression correlation occurs.

Currently, exports are not included as a model prediction parameter. Should exports be included as a model prediction parameter (for smolt production)?

I think exports should be included but if they are brought in consistent with the other factors in the model, then I suspect not much good will result. The author seems bent on only including statistical significant relationships. Yet, some of the relationships included are questionable and seem to me to be at the same level of confidence of other relationships dismissed as “not significant enough” to be included. Exports seem to be an example of this. To me, I would include exports and density-dependence and ocean harvest, given the way the model is currently being used.

6. Are the methods used in the Model (including Model Report) relating to statistical evaluations and/or model logic justification sufficiently documented? If not, what improvements can be made to improve documentation?

The logic of model development is well presented. There are too many instances where it seems the author deems one relationship “not significant enough” to include while including other relationships that look to me to be similarly weak.

7. What improvements to statistical use and application can be made to enhance model prediction performance reliability?

More sophisticated statistical methods can be applied to some of the critical datasets. There are well established alterative methods to simple linear regression, such as robust regression and regression trees. The issue of serial correlation due to the data being mostly time series is ignored, perhaps some regression-like methods that explicitly account for time series data could be investigated.

8. There is substantial disagreement amongst scientists regarding the issue of density dependent mortality and its influence upon SJR salmon abundance (e.g. fall spawner abundance and spawning habitat availability: aka redd superimposition). In the absence of flow the relationship between spawner abundance and stock recruit appears to show density dependence (i.e. Beaverton-Holt or other density dependent type relationship). However when flow is included with spawner abundance, in the form of a multiple-regression using spawner abundance and spring flow regressed against adult recruits, a significant correlation exists suggesting that density dependence does not explain the variation in SJR adult salmon escapement abundance. How can this issue be resolved with data to date, or in the future if data insufficiency exists currently?

I think the analyses dismissing density-dependence in the report is weak. I look at Figures 28 and 29 and I see the potential for density-dependence. The author states “Both the increased fry density with increased spawner density, and increased cohort abundance with increased spawner abundance, are contrary with the density-dependent hypothesis (page 16).” Either I mis-understand this statement or it is wrong. Both increasing with increasing spawners is consistent with density-dependence, it is the rate of their increase that tells you whether density-dependence is strong or weak. I think the presently available data are sufficient to resolve the density-dependence issue, or at least sufficient to specify a relationship and see if it affects the predictions and conclusions. Some of the scenarios involve higher salmon abundance which could trigger density-dependence.

9. How can the statistical relationships between flow and fish survival and/or fish production be improved?

One way to improve them is to use more sophisticated statistical methods than just linear regression. Even including regression diagnostics about outliers and influence points, patterns in residuals, and bootstrapping would help the rigor of the results. Perhaps some of these datasets are better suited for other methods than linear regression, such as robust regression, survival analysis (for the mortality data), regression trees, and time series methods...

**Miscellaneous comments:**

This modeling approach has the advantage of appearing to be based on statistical relationships and thus might appear to be more defensible. This often turns out to be a false advantage. People argue about outliers, whether the relationships included should be linear or not, and about the relationships not included. If you want to use the model to suggest that more flow (within reason and practicable amounts) and a longer, delayed time window would help the salmon, then I agree with the conclusions. If you want to specify the flow to within 1000 cfs and the time window within days or weeks, then I have problems with the current formulation of the model.

(1) page 10.—the argument for focus on smolt is confusing. For example, the author says more fry means higher escapement but also higher smolt abundance, as if this argues for smolts being more important than fry. I would think more fry would mean more smolts. If there is evidence of low DO having an effect, then it should be summarize and cited. So smolt and escapement are correlated, does that mean fry are unimportant? Some clarification is needed here.

(2) page 11.—I agree that the Delta may not be a good place anymore for salmon. But the fact they grew fast in the ocean does not mean the Delta is not helpful. There are also predation issues.

(3) page 12.—Figures 9 and 10 are interesting. I think they mean that VAMP is dealing with a small amount of water.

(4) page 13.—the model really does not allow for confidence intervals to be calculated.

(5) page 14.—the argument about the unimportance of exports is difficult to follow and seems convoluted. Evidence of this is footnote 19. I am not convinced that the author has shown that it is flow and not exports. The data by Pat Brandes has been the subject of much discussion and to use the way the author is using the data is risky. I am not convinced that the author has explained away Figure 18, especially if the model is required to generate absolute

predictions of escapement and be able to do this for difference of 1000's of cfs of flow. I think changes in operations confound the entire picture.

(6) page 15.—I disagree that ocean harvest can be treated as a constant. The relationship dismissed in Figure 26 is as strong as or stronger than some of the relationships included in the model (for example figure 36).

(7) page 16.—I found the discussion about density-dependence to be confusing and I think misguided. I do not follow the argument about how figure 28 is not showing density-dependence?

(8) page 17.—Figure 29 is presented as the evidence that Vernalis flow is important; yet, the relationship shown in the figure is very weak evidence. I would guess that the relationships depends on the two extreme points. The rest of the data is a cloud.

(9) page 19.—Figures 34 and 36 are critical to the model and pretty much dictates how important flow is. These linear relationships are suspect.

(10) page 22.—The model calibration is poorly documented. Which parameter were varied and how? How close are their final value to the values estimated from the data? Why not use an optimization (minimization) method.

(11) page 23.—There are methods for propagating uncertainty through these models.

(12) page 25.—The results are nicely presented.

(13) page 26.—Why not optimize the model to obtain a full exploration of the trade-offs between water usage and escapement enhancement?

(14) page 27.—The model does not provide evidence that spring flow is important. The model was built under that assumption. Be clear that the data were analyzed and the author concluded that flow was important. It is not a model result.